

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2002, 22:21:02 ; Search time 26 seconds

(without alignments)

660.431 Million cell updates/sec

Title: US-09-708-724a-2

Perfect score: 2187

Sequence: 1 MGFWSVVLVCGMKQLGQAL.....LLAVTREGLEIRISKRRAE 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135.5	6.2	563	1	ARX_MOUSE
2	131	6.0	2090	1	HFC1_MESAU
3	121.5	5.6	1395	1	IF4G_HUMAN
4	119	5.4	741	1	GTSE_MOUSE
5	117.5	5.4	1402	1	IF4G_RABIT
6	116	5.3	3149	1	TEGU_EBV
7	112.5	5.1	584	1	BCAS_HUMAN
8	112.5	5.1	1386	1	ZAP3_MOUSE
9	111.5	5.1	792	1	PCAP_MOUSE
10	111.5	5.1	2220	1	CABI_HUMAN
11	111	5.1	728	1	VIV_ORYSA
12	110.5	5.1	1146	1	CCAS_RAT
13	110.5	5.1	1290	1	PER1_HUMAN
14	110	5.0	817	1	VRP1_YEAST
15	110	5.0	2142	1	BAT2_HUMAN
16	109.5	5.0	1575	1	SYJ1_HUMAN
17	109	5.0	883	1	PCCB_RAT
18	109	5.0	1004	1	SAL2_MOUSE
19	107.5	4.9	599	1	RF2P_DROSI
20	107	4.9	1260	1	ALSI_CANAL
21	107	4.9	1388	1	CAIE_HUMAN
22	106.5	4.9	365	1	VNSA_HPBDC
23	106	4.8	620	1	EXTN_TORAC
24	106	4.8	633	1	LA17_YEAST
25	106	4.8	838	1	TAC3_HUMAN
26	106	4.8	1233	1	NME3_HUMAN
27	106	4.8	1822	1	ZAP3_HUMAN
28	105.5	4.8	626	1	GPBA_HUMAN
29	105.5	4.8	743	1	TLE2_HUMAN
30	105.5	4.8	1210	1	AR4_HUMAN
31	105.5	4.8	2715	1	TRX2_HUMAN
32	105	4.8	671	1	CHS5_YEAST
33	105	4.8	751	1	TAU_RAT

RESULT 1				
ARX_MOUSE	ID	ARX_MOUSE	STANDARD;	PRT; 563 AA.
AC	O35085;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Homeobox protein ARX.			
GN	ARX.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97398450; PubMed=9256348;			
RA	Miura H., Yanazawa M., Kato K., Kitamura K.;			
RT	"Expression of a novel aristless related homeobox gene 'Arx' in the			
RT	vertebrate telencephalon, diencephalon and floor plate.";			
RL	Mech. Dev. 65:99-109(1997).			
CC	-!- SUBCELLULAR LOCATION: Nuclear (By similarity).			
CC	-!- SIMILARITY: BELONGS TO THE PAIRED HOMEBOX FAMILY.			
CC	-!- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.			
CC	-!- SIMILARITY: CONTAINS 1 OAR DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB006103; BAA28284.1; -			
DR	HSP; P06601; IFJL.			
DR	MGI; 1097716; Arx.			
DR	InterPro; IPR000047; HTH_repressr.			
DR	InterPro; IPR003654; Homeo_OAR.			
DR	InterPro; IPR001356; Homeobox.			
DR	Pfam; PF00046; homeobox; 1.			
DR	PRINTS; PR00024; HOMEBOX.			
DR	PRINTS; PR00031; HTHREPRESSR.			
DR	PRODom; PD000010; Homeobox; 1.			
DR	SMART; SM00389; Hox; 1.			
DR	PROSITE; PS00027; HOMEBOX_1; 1.			
DR	PROSITE; PS50071; HOMEBOX_2; 1.			
DR	PROSITE; PS50803; OAR; 1.			
DR	Homeobox; DNA-binding; Nuclear protein.			
KW	Homeobox; 330 389			
FT	DNA_BIND 330 389			
FT	DOMAIN 531 544			
FT	DOMAIN 100 116			
FT	DOMAIN 127 132			
FT	DOMAIN 143 154			
FT	DOMAIN 183 186			
FT	DOMAIN 231 241			
FT	DOMAIN 274 285			


```
QY 329 LREVEWPGRGH-----MAATCKKLV-----EGDRTMSLAAAPVREAPPPPTG 372
Db 1572 LQASSTPLMGSESVLSASPTPELQAAVDLSTGDPSSGQEPSSAVVATVVQPPPTQ 1631
QY 373 AS-SEPSVPALPGADPQPSAEILL-----LAVT 399
Db 1632 SEVDLSLPQELMAEQAAGTTLMTVGLTPELAVT 1667

RESULT 3
ID IF4G_HUMAN STANDARD; PRT; 1395 AA.
AC Q04637;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eukaryotic translation initiation factor 4 gamma (eIF-4 gamma) (eIF-4G) (eIF4G) (P220).
DE 4G) (eIF4G) (P220).
DE
DE
GN EIF4G1 OR EIF4G.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93054654; PubMed=1429670;
RA Yan R., Rychlik W., Etchison D., Rhoads R.E.;
RT "Amino acid sequence of the human protein synthesis initiation factor
RT eIF-4 gamma.";
RL J. Biol. Chem. 267:23226-23231(1992).
RN [2]
RP SEQUENCE OF 408-525 FROM N.A.
RX MEDLINE=95379845; PubMed=7651417;
RA Mader S., Lee H., Pause A., Sonenberg N.;
RT "The translation initiation factor eIF-4E binds to a common motif
RT shared by the translation factor eIF-4 gamma and the translational
RT repressors 4E-binding proteins.";
RL Mol. Cell. Biol. 15:4990-4997(1995).
CC
CC
CC -!- FUNCTION: COMPONENT OF THE PROTEIN COMPLEX EIF-4, WHICH IS
CC INVOLVED IN THE RECOGNITION OF THE MRNA CAP, ATP-DEPENDENT
CC UNWINDING OF 5'-TERMINAL SECONDARY STRUCTURE AND RECRUITMENT OF
CC MRNA TO THE RIBOSOME.
CC
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES IN VIVO.
CC
CC -!- SIMILARITY: THE C-TERMINAL REGION IS SIMILAR TO THE N-TERMINAL
CC REGION OF WHEAT GERMINAL INITIATION FACTOR (ISO)4F SUBUNIT P82.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12686; BAA02185.1; -
DR Genbank; HGNC:3296; EIF4G1.
DR MIM; 600495; -
DR InterPro; IPR003890; IF_eIF4G.
DR InterPro; IPR003891; IF_eIF4G_MA3.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF02020; W2; 1.
DR Pfam; PF02847; MA3; 1.
DR Pfam; PF02854; MIF4G; 1.
DR SMART; SM00544; MA3; 1.
DR SMART; SM00543; MIF4G; 1.
DR SMART; SM00515; eIF5C; 1.
KW Initiation factor; Protein biosynthesis; Phosphorylation;
FT DOMAIN 258 271 POLY-GLU.
FT DOMAIN 305 308 POLY-ALA.
FT CONFLICT 497 500 PRGP -> ARGAQ (IN REF. 1).
```

```
SQ SEQUENCE 1395 AA; 153360 MW; C712A09D270DF2A8 CRC64;
Query Match 5.6%; Score 121.5; DB 1; Length 1395;
Best Local Similarity 24.1%; Pred. No. 1.4;
Matches 60; Conservative 28; Mismatches 92; Indels 69; Gaps 11;

QY 167 NKGGRSCQNPALPQDQSPSQGNATTSVTRDNYHLLTTEEEFGVWSQMKWHSONKSGGVP 226
Db 40 SQGAIADRPGLPGPEHSPS-----ESQSPSPSPPT 70

QY 227 VRGPTQEPCESEQLIKESFVPPT-----TPKENNKQRE-DENWRLPPPPVAETPVSP 279
Db 71 SPSPVLEPGSEPNLAVLSIPGDTMTIOMSVSEETPSRETGETPYRUSPEP---TPLAEP 127

QY 280 SVTETETPLQIPRTATITAGEPLGHCCTTISPAFVHSLNKRKQLELLREVEWPGRGH 339
Db 128 -LLEVEVTLSPVSESEFSSPLQAPT-----PLASHTV-----EIHEP-NGM 168

QY 340 MAATCKKLVQEGQDRTMSLAAAPVREAPPP--PT-----GASSEPSVPALPGADPQR 389
Db 169 VPSEDLPEVESSPELAPPACP-SESPVPIAPTAQPELLNGAPSPPAVDLSVPVSEDEE 227

QY 390 SAEILLIAY 398
Db 228 QAKEVTASV 236

RESULT 4
ID GTSE_MOUSE STANDARD; PRT; 741 AA.
AC Q8R080; O89015; Q9CGSG;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE G2 and S phase expressed protein 1 (gtse-1) (B99 protein).
DE
DE
GN GTSE1 OR B99.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98393542; PubMed=9724637;
RA Utrera R., Collavin L., Lazarevic D., Delia D., Schneider C.;
RT "A novel p53-inducible gene coding for a microtubule-localized protein
RT with G2-phase-specific expression.";
RL EMO J. 17:5015-5025(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-130 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
```



```
Query Match          5.4%; Score 117.5; DB 1; Length 1402;
Best Local Similarity 26.4%; Pred. No. 2.7;
Matches 63; Conservative 18; Mismatches 89; Indels 69; Gaps 11;

QY 159 SGGNIINTKGRSCQNALPSPDOSPSGNATTSTVRDNYHLLTTEERGVWSQSMKWHQ 218
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 40 SQGAIL-----GGR-----PGLPGPEHSPS-----ESQ 62
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 219 NKSGSVVRGPTORPCSESOILKESFVPTTPKENNKQEREDENWRLLPPPPVAETVPV- 277
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 PSSPSPTPSPPILEPGSEPNLAVLS-LPGDTWISGMLOMPVEE-----PAPISREAGEPY 117
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 278 --SPSVTEIETPLQRIPTATAGBPLGCHCTFTTISPAFVSHVLNKRKRQLELLLEVEWMP 335
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 118 CLSPEPTPLAEPILEVEVTL-----KVPVSEFFSSPIQVLTPLASHK-----MEIHEP 167
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 336 GRGHMAATCKLQVGGQDRTMSLAAAPVREAPPPPTGASSESVPALPGADPQRSALL 394
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 168 NG-----VVPSEDLPEVESSP-ELAPPPPPACLSSESPVPIAPTQPE-----ELL 212
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
TEGU_EBV
ID TEGU_EBV STANDARD; PRT; 3149 AA.
AC P03186;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein.
GN BPLF1.
OS Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrall B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
CC -1- FUNCTION: TEGUMENT PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01555; CAA24839.1; -.
CC PIR; A03747; QQBEB8.
CC PIR; S32993; S32993.
CC
CC SEQUENCE 3149 AA; 337954 MW; 3DD0C576587313D8 CRC64;

Query Match          5.3%; Score 116; DB 1; Length 3149;
Best Local Similarity 24.6%; Pred. No. 8.7;
Matches 81; Conservative 27; Mismatches 111; Indels 110; Gaps 18;

QY 103 PTTAPPLTSGNVTAEDEAIRAFVYAAASAAAEAAHWRHLVLLSGQIHEPIGSGGN 162
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 340 PASAPASAAAPASAPASAPASAPASA-APASAPA-----SPPFLFIPDGLGH 390
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 163 INTNKGGRSCQNALPSPDOSPSGNATTSTVRDNYHLLTTEERGVWSQSMKWHQSN 219
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 -----TPGVAPAPSTPPRASSGAAPQTPKRKK-----GLGRDSP--HKKP 427
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 220 KSGSVVVRGPTOEPCSESOILKESFVPTTPKENNKQEREDENWRLLP-----PP 270
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
FT DOMAIN 180 199 POLY-GLN.
FT DOMAIN 211 222 POLY-GLN.
FT DOMAIN 230 242 POLY-GLN.
FT DOMAIN 246 259 POLY-GLN.
FT DOMAIN 453 460 POLY-PRO.
FT DOMAIN 603 609 POLY-PRO.
SQ SEQUENCE 792 AA; 87061 MW; 71C879DCD3925E62 CRC64;

Query Match 5.1%; Score 111.5; DB 1; Length 792;
Best Local Similarity 22.8%; Pred No 3.4;
Matches 87; Conservative 49; Mismatches 145; Indels 101; Gaps 20;

Oy 33 GKRCPCFGAQLMTCQNTPLSPVSHRSPGNAAVSVTGDCGCHLPTEBEFGVLVQSMKCD 92
Db 301 QAQOSPI--AQN---QPPQIPQSQSQPLVSRAQALPGPMLYAAQQO-----LK 344
Oy 93 TVRKGVLOGTTPAPLMTSGNVTAEDETEAIRAFVYAVAAASAAHWRHLVLLSG- 151
Db 345 FVRAPMVVQOPQOVQOVQOVQO-----AAVQAQSAQ-----MVAPGV 387
Oy 152 -QIHEPIGSGNIINTNKGGRSCONPALPSPDQSPGNATTSVTRDNYHLLTEEEFG-- 207
Db 388 QMIAEALAQGMHVRAPPTSTWS-AGPSSSISLGGOPTTQVQSSUTMLSSPSPGQOV 446
Oy 208 VWSQSMKWHQO-NKSGGSPV---VRGPTQPCSESQILKESFPVPTTPKNNK--QERE 260
Db 447 QTPQSMPPPPQSPQSPQSPNSVSSGPASP-----SSFLPSPSPQSPQSVTART 498
Oy 261 DENWRLPPPPVAETPVPSPV-----TEIETPLQRIPTATYAGPLGHCFTTISPA 312
Db 499 PONSFPSPGFLPNTPVNPSVMSFAGSSQAEEQQYLDKQLSKYI-BPL----- 547
Oy 313 FVHSLVNL-----RKROL-----ELLREVPEPGRGHMAATC--CKLOVEGQDRTMSLAA 361
Db 548 --RMWINKIDNEDKKDLKMSLLDITDPSKRCPLTKLQKCEIALEKLNDAW--- 602
Oy 362 PVREAPPPPTGASSEPSVPALP 383
Db 603 ---PTPPPP-----PVLPL 612

RESULT 10
CABI_HUMAN STANDARD; PRT; 2220 AA.
AC Q9Y6J0; Q9Y460;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcineurin-binding protein Cabin 1 (Calcineurin inhibitor) (CAIN).
GN CABI1 OR KIA0330.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98318216; PubMed=9655484;
RA Sun L., Yoon H.-D., Loh C., Stolor M., He W., Liu J.O.;
RT "Cabin 1, a negative regulator for calcineurin signaling in T
RL lymphocytes.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Dawson E.,
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
```

```
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grahame D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Swann R.M.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [3]
RP SEQUENCE OF 319-2220 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC -!- FUNCTION: IT MAY SERVE AS A NEGATIVE REGULATOR OF T CELL RECEPTOR
CC (TCR) SIGNALING VIA INHIBITION OF CALCINEURIN. INTERACTS WITH AND
CC INHIBITS CALCINEURIN-MEDIATED SIGNAL TRANSDUCTION. CABIN 1 IS
CC SPECIFIC FOR THE ACTIVATED FORM OF CALCINEURIN. THEIR INTERACTION
CC IS DEPENDENT ON BOTH PKC AND CALCIUM SIGNALS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT TISSUES.
CC -!- PTM: ACTIVATED THROUGH PKC-MEDIATED HYPERPHOSPHORYLATION.
CC -!- SIMILARITY: CONTAINS 6 TPR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF072441; AAD40846.1; -;
CC EMBL; AP000352; -; NOT_ANNOTATED_CDS.
CC EMBL; AP000353; -; NOT_ANNOTATED_CDS.
CC EMBL; Z92546; CAB62954.1; -;
CC EMBL; AB002328; BAA20788.1; -;
CC MIM; 604251; -;
CC InterPro; IPR001440; TPR.
CC Pfam; PF00515; TPR; 4.
CC SMART; SM00028; TPR; 3.
KW phosphorylation; Repeat; TPR repeat.
```

```

FT REPEAT 36 69 TPR 1.
FT REPEAT 90 123 TPR 2.
FT REPEAT 125 157 TPR 3.
FT REPEAT 615 648 TPR 4.
FT REPEAT 1055 1088 TPR 5.
FT REPEAT 1106 1139 TPR 6.
FT DOMAIN 2116 2153 CALCINEURIN BINDING (BY SIMILARITY).
SQ SEQUENCE 2220 AA; 246350 MW; BA4AD1741056C233 CRC64;

Query Match
Best Local Similarity 5.1%; Score 111.5; DB 1; Length 2220;
Matches 104; Conservative 39; Mismatches 171; Indels 147; Gaps 21;

QY 10 VCGMKQLGALQASVLSIITENQGRKPCFCGAQNLMTQNPLPSVSHRSPPGNAAVSV 69
Db 1678 VCGLP--GARMTTDVS-----HKASPEDQOGL---PQPKPLADSGPGPEPGK 1724
QY 70 TGGDCHLPTREFFGVLMQSKCDTVRIKGVLO-----GPTAPPLMTSEGNVTAEDTE 122
Db 1725 VGLLNHRPVAMDAG-----DSADQSGERKDKESPRAGPT--EPMDDTSEATVCHSDLE 1774
QY 123 EAIRAFVYVAARASAAEA---WHRHLVLLSQIHEPIGSGNININTKGRSQNPAL 178
Db 1775 RTPPLPGRPADRPESRPTLSLELSISARQOFTPL-----TPAQ 1817
QY 179 PSPDQSPSGNATTSVTRONYH-----LITEEEFG-----VWSQSMKWH 216
Db 1818 PAPAPAP---ATTGTRAGHPPEPLSLRSRKRKLLLEDTESKTLLLDAYRVWQOGK-- 1872
QY 217 SONKGGSVVPGVPQPCSESQIL-----KESFVPPPT 249
Db 1873 -----GVAYDLGRVERIMSETYMLIKQVDEAALEQAVKFCQVHGAQAQASGDTPT 1926
QY 250 TPKENNKQEREDENRLPPPPVAETVPSPSVTEIETPLQRIPTATATAGEPLGHTFTI 309
Db 1927 TPK--HPKDSREN--FPVTVVTPADVPADSV--QRPDAHTKPRPALAAATTI-----ITC 1979
QY 310 SPAFVHSLNKKRQLELLREVEWPG-----RGHMAATCCCKLOVEGQD--RTMSLAAP 362
Db 1980 PFSASASTLDQSKD-----PGPRPHRPEATPSMASLGPGEDELARVAEGTSFP 2028
QY 363 VREAPPPPTGASSEFSPVLPAL-----GADPQRSAAE 392
Db 2029 PQEPRHSPQVKMAPTSSPAEPHCWFAEAALGTGAETCSQE 2069

RESULT 11
VIV_ORYSA STANDARD; PRT: 728 AA.
AC P37398;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Viviparous protein homolog.
GN VPI.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94250843; PubMed=8193305;
RA Hattori T., Terada T., Hamauna S.;
RT "Sequence and functional analyses of the rice gene homologous to the
maize Vpi.";
RL Plant Mol. Biol. 24:805-810(1994).
CC -!- FUNCTION: COULD PARTICIPATE IN ABSICISIC ACID-REGULATED GENE
EXPRESSION DURING SEED DEVELOPMENT.
CC -!- SIMILARITY: CONTAINS 1 TF-B3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16640; BAA04066.1; -.
DR PIR; S48899; S48899.
DR TRANSFAC; T04785; -.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 1.
KW Developmental protein; Transcription regulation; DNA-binding;
KW Activator.
FT DOMAIN 500 685 TF-B3.
SQ SEQUENCE 728 AA; 75982 MW; F66882079F5FE428 CRC64;

Query Match
Best Local Similarity 5.1%; Score 111; DB 1; Length 728;
Matches 67; Conservative 29; Mismatches 114; Indels 92; Gaps 15;

QY 170 GRSCONPALPSD-----QSPSGN-----ATTSVTRDNYHLLTEEFVWSQSMKWHSQ 218
Db 257 GAASDQPLPSPCANPGVEFFSGQEMGSAATS-----WMPYQAFTPP 300
QY 219 NKSGSVVRG-----PTQPCSESQILKES--FVPTTPKENNKQEREDENRLP----- 267
Db 301 AAYGDAMYPGAAGPFPFQSCSKSVVSSQFPPTAAAAGDMHASGGGNWAPQOFA 360
QY 268 PPVVAETVPSPSVTEIETPLQRIPTATATAGE-PLGHTFT-----ISPAFVHSLNK- 320
Db 361 PPVSVST--SSYTMPSPVVP-----PFTAGFPGQYSGGHAMCSPRLAGVPESTKEARKR 414
QY 321 --RKQLELLREVEWPGRGHMAATCCCKLOVEGQRTMSLAAA----- 361
Db 415 MARQRLSCLOQ-----RSQQLNLQIHSHGHQEPSRAHSAVPYTPSSAGCSWGI 468
QY 362 -----PVREAPPPPTGASSEFS-VPALPGADPQPSAEALLLVAVTREGLEIRISR 410
Db 469 WPPAAQIIQNPLSNPNPPATSKQPKSPKPKPQAAA-----TAGAESLQSTASE 523
QY 411 KR 412
Db 524 KR 525

RESULT 12
CCAS_RAT STANDARD; PRT: 1146 AA.
AC Q02485; Q01553; P70484; Q62817;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1S subunit (Calcium
channel, L type, alpha-1 polypeptide, isoform 3, skeletal muscle)
DE (ROB1) (Fragment).
GN CACNA1S OR CACNA1 OR CACN1 OR CCHL1A3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93122775; PubMed=1335956;
RA Chin H., Krall M., Kim H.-L., Kozak C.A., Mock B.A.;
RT "The gene for the alpha-1 subunit of the skeletal muscle
RT dihydropyridine-sensitive calcium channel (Cch1a3) maps to mouse
chromosome 1.";
RL Genomics 14:1089-1091(1992).
RN [2]
RP SEQUENCE OF 597-691 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93066265; PubMed=1279681;
RA Yu A.S.L., Hebert S.C., Brenner B.M., Lytton J.;

```


"Molecular characterization and nephron distribution of a family of transcripts encoding the pore-forming subunit of Ca²⁺ channels in the kidney.";
Proc. Natl. Acad. Sci. U.S.A. 89:10494-10498(1992).
[3]
SEQUENCE OF 359-658 FROM N.A. (ISOFORM ROB1).
TISSUE-Osteosarcoma;
MEDLINE=96074617; PubMed=7479909;
Barry E.L.R., Gesek F.A., Froehner S.C., Friedman P.A.;
"Multiple calcium channel transcripts in rat osteosarcoma cells: selective activation of alpha 1D isoform by parathyroid hormone.";
Proc. Natl. Acad. Sci. U.S.A. 92:10914-10918(1995).
[4]

PHOSPHORYLATION BY CAPK.

MEDLINE=91065881; PubMed=2174428;
Lai Y., Seagar M.J., Takahashi M., Catterall W.A.;
"Cyclic AMP-dependent phosphorylation of two size forms of alpha 1 subunits of L-type calcium channels in rat skeletal muscle cells.";
J. Biol. Chem. 265:20839-20848(1990).
-!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION, CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1S GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE) CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP).
PHENYLALANINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-1I1A (OMEGA-AGA-1I1A). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA). CALCIUM CHANNELS CONTAINING THE ALPHA-1S SUBUNIT PLAY AN IMPORTANT ROLE IN EXCITATION-CONTRACTION COUPLING IN SKELETAL MUSCLE.
-!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY. AN ADDITIONAL GAMMA SUBUNIT IS PRESENT ONLY IN SKELETAL MUSCLE L-TYPE CHANNEL.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: SKELETAL MUSCLE SPECIFIC.
-!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
-!- DOMAIN: THE LOOP BETWEEN REPEATS II AND III INTERACTS WITH THE RYANODINE RECEPTOR, AND IS THEREFORE IMPORTANT FOR CALCIUM RELEASE FROM THE ENDOPLASMIC RETICULUM NECESSARY FOR MUSCLE CONTRACTION.
-!- PTM: PHOSPHORYLATION BY CAPK STIMULATES THE CALCIUM CHANNEL FUNCTION (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L04684; AAA00844.1; -;
EMBL; M99220; AAA00894.1; -;
EMBL; U31816; AAA89158.1; -;
InterPro; IPR001682; Ca/Na_pore.
InterPro; IPR002111; Cat_channel_Trlp.
InterPro; IPR000636; M-channel_nlg.
Pfam; PF00520; Ion_trans; 2.
Ionic channel; Transmembrane; Ion transport; Voltage-gated channel; Calcium channel; Glycoprotein; Repeat; Multigene family;

KW	Calcium-binding; Phosphorylation; Alternative splicing.	1
FT	NON_TER	1
FT	REPEAT	355
FT	REPEAT	392
FT	DOMAIN	<1
FT	TRANSMEM	87
FT	DOMAIN	106
FT	TRANSMEM	122
FT	DOMAIN	142
FT	TRANSMEM	154
FT	DOMAIN	173
FT	TRANSMEM	180
FT	DOMAIN	199
FT	TRANSMEM	218
FT	DOMAIN	238
FT	TRANSMEM	328
FT	DOMAIN	353
FT	TRANSMEM	406
FT	DOMAIN	425
FT	TRANSMEM	440
FT	DOMAIN	460
FT	TRANSMEM	468
FT	DOMAIN	487
FT	TRANSMEM	519
FT	DOMAIN	538
FT	TRANSMEM	557
FT	DOMAIN	577
FT	TRANSMEM	644
FT	DOMAIN	669
FT	SITE	301
FT	SITE	610
FT	BINDING	275
FT	BINDING	624
FT	BINDING	636
FT	MOD_RES	679
FT	CA_BIND	697
FT	CARBOHYD	428
FT	VARSPLIC	491
FT	CONFLICT	548
FT	CONFLICT	610
FT	SEQUENCE	1146 AA; 130143 MW; 4BBE944261BE1A95 CRC64;

Query Match

Best Local Similarity	22.8%;	Pred No. 6.1;	Length 1146;
Matches	82;	Conservative	45; Mismatches 139; Indels 93; Gaps 17;
Qy	79	EEFGVLYQSMKCDTVRIKGVQLQ--GPTTAPPLMTS-EGNVTAEDETEAIRAFVAVAAA	135
Db	828	EYVG---YRPKDTVQIQAGLRTIEEAAPEIHRASIGDLTAE--EELERAMVEA----	878
Qy	136	SAEAHWRHLVLLSGQIHEPIGSGGNININTKGRSC-----QN 175	
Db	879	AMEEGIFRTGGLFGVDNLFLENTSLPPVMANORLQFAEMEMEELSPVLEDFPON 937	
Qy	176	PALPSPDQSPSNATSVTRDNYHLLTEEEFGVWSQSMKHSQ--NKSGGSPVVRGPTQE 233	
Db	938	PGTHPLARNTNANNAVYGNSSHRNSPVF---SSINERELLEAGRPVTRGPTSQ 993	
Qy	234	PCSESQILKESFPVPTTPKENNKQREDENWRLPPPPVAETPVP-----277	
Db	994	PCSVSGVNSRSHV-----DKLERQMSQRMKPGVPPSPCOLSQEKKHPVHEBKGPR 1045	
Qy	278	--SPSVTEIETPLQRIPTATTAGEPLGH-----CTFTISPAFVHSVINKRKRQLELL 329	
Db	1046	SWSTETSDSESPERVPNRNS-----AHKCTAPATTMLIQEALVRGGLDLSAADANFVM 1098	
Qy	330	REVENPGRGHMAATCCKLQVEQDRTMSLAAAPV--REAPP--PTTCASESPSPALPGA 385	
Db	1099	-----ATQALADACOMEPE-----EVEVAATELLKRPKGGPCGSLPKV--LPWA 1145	

RESULT 13
 PERL_HUMAN STANDARD; PRT; 1290 AA.
 AC O15534;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Period circadian protein 1 (Circadian pacemaker protein Rigiui) (hPER).
 GN PERL OR PER OR RIGUI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Heart;
 RX MEDLINE=97462301; PubMed=9323128;
 RA Sun Z.S., Albrecht U., Zhuchenko O., Bailey J., Eichele G., Lee C.C.;
 RT "Rigiui, a putative mammalian ortholog of the Drosophila period gene.";
 RL Cell 90:1003-1011(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97472418; PubMed=9333243;
 RA Tei H., Okamura H., Shigeyoshi Y., Fukuhara C., Ozawa R., Hirose M.,
 RA Sakaki Y.;
 RT "Circadian oscillation of a mammalian homologue of the Drosophila
 period gene.";
 RL Nature 389:512-516(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Hida A., Sakaki Y., Tei H.;
 RX PubMed=10940553;
 RA Taruscio D., Zorraqi G.K., Falchi M., Iosi F., Paradisi S.,
 RA Di Fiore B., Lavia P., Falbo V.;
 RT "The human Perl gene: genomic organization and promoter analysis of
 the first human orthologue of the Drosophila period gene.";
 RL Gene 253:161-170(2000).
 CC -!- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
 FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
 LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
 TRANSCRIPTIONAL INHIBITION. SEEMS TO BE THE PACEMAKER COMPONENT
 WHICH RESPONDS TO LIGHT AND MEDIATES PHOTIC ENTRAINMENT. IN THE
 SUPRACHIASMATIC NUCLEUS (SCN), IT BEHAVES LIKE A DAY-TYPE
 OSCILLATOR, WITH MAXIMUM EXPRESSION DURING THE LIGHT PERIOD.
 OSCILLATIONS ARE MAINTAINED UNDER CONSTANT DARKNESS AND ARE
 RESPONSIVE TO CHANGES OF THE LIGHT/DARK CYCLES. THERE IS A 4 HOUR
 TIME DELAY BETWEEN PER1 AND PER2 OSCILLATIONS. THE EXPRESSION
 RHYTHMS APPEAR TO ORIGINATE FROM RETINA (BY SIMILARITY).
 CC -!- SUBUNIT: FORMS A HETERODIMER, PROBABLY WITH CLOCK.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; RIGUI 4.7 (SHOWN HERE),
 RIGUI 3.0 AND RIGUI 6.6/TRUNCATED; ARE PRODUCED BY ALTERNATIVE
 SPLICING.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. FOUND IN HEART, BRAIN,
 PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, PANCREAS, AND AT LOW LEVEL
 IN THE KIDNEY.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS.
 CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
 CC EMBL; AF022991; AAC51765.1; -
 CC EMBL; AB002107; BAA22633.1; -
 CC EMBL; AB030817; BAA94085.1; -
 CC EMBL; AF102137; AAF15544.1; -
 CC Genew; HGNC:8845; PER1.
 CC MIM; 602260; -
 CC InterPro; IPR001610; PAS.
 CC InterPro; IPR000014; PAS_domain.
 CC Pfam; PF00989; PAS; 1.
 CC SMART; SM00086; PAS; 1.
 CC SMART; SM00091; PAS; 2.
 CC PROSITE; PS01112; PAS; 1.
 CC Transcription regulation; Nuclear protein; Repeat; Biological rhythms;
 KW Alternative splicing.
 FT DOMAIN 136 172 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 209 275 PAS 1.
 FT DOMAIN 346 416 PAS 2.
 FT DOMAIN 425 465 PAC.
 FT DOMAIN 66 79 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 69 75 POLY-SER.
 FT DOMAIN 653 656 POLY-SER.
 FT DOMAIN 1010 1013 POLY-PRO.
 FT DOMAIN 1269 1273 POLY-GLU.
 FT DOMAIN 1276 1279 POLY-SER.
 SQ SEQUENCE 1290 AA; 136237 MW; 24B53042869A4562 CRC64;
 Query Match 5.1%; Score 110.5; DB 1; Length 1290;
 Best Local Similarity 20.1%; Pred. No. 7;
 Matches 91; Conservative 42; Mismatches 162; Indels 157; Gaps 20;
 QY 36 RCPFCGA--QNLMTQNTPLPSVSHRSPGNAASVTGGD----- 73
 DB 576 RLPAATGTFKAKALPCQSPD-PELEAGSAPVQAPLALVPEEAERKEASSCYQINCLDSI 634
 QY 74 -----CHLPTEEFGLVQSMKC-----DTVRKGVLCQPTTAPPLMTSE 113
 DB 635 LRYLESCNLP-----TTKRKCASSSYTTSSASDDDRQRTGPVSVGTTKDPSPSAALS 687
 QY 114 GNVTAEDTEATRAFYVAASAAENWHRLVLLSQ-----IHEPIS----GCN 162
 DB 688 GEGATRKRPVVGGLSLALANKAES-----VVSVCSCSFSSSTIVH--VGDKKRPESD 740
 QY 163 IINTNKGGRSCNPA---LPS-----PDQSPGNATTSVTRDNYHLTTEE----- 205
 DB 741 IIMMEDPLCLAGPAPSPAPSPVAPDPADPAPVGLTKAVLSLHTKEQEAFLSRFRD 800
 QY 206 -----FGVWSQSMKWSKSGSVVPGVGTQOE-PCSESOIL 241
 DB 801 LGLRLGLDSSSTAPSGALGERGCHGPPAPPSRRHRCRKAARSRHHQNPRAEAPCYVSH-- 858
 QY 242 KESFVPPPTPKENNKOEDENNRLPPP---PVAETPVPSVTEIETPLQRIPTATI 297
 DB 859 -PSPVPPSTP-----WPTTPATPPAVVQVPLPVFS-----PRGGP- 895
 QY 298 AGEPLGHCTTISPAFVSHVNLKRRKQLELLREVEWPGRGHMAATCCQLQVEGQDRTMS 357
 DB 896 --QPLPAPPTSPVPAFAFPAPL--VTPMVALVLPNLYLFTPTSPSSYPYQALQTPAEG----- 945
 QY 358 LAAAPVREAPPPTGASSEPSVPALPGADPQR 389
 DB 946 -----PPTPASHSPSPSLPALPPSPPHR 968
 RESULT 14
 VRPL_YEAST STANDARD; PRT; 817 AA.
 ID VRPL_YEAST
 AC P37370; Q06133;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Verprolin.

KW	Repeat.	519	524	POLY-PRO.
FT	DOMAIN	636	657	GLN-RICH.
FT	DOMAIN	684	688	POLY-PRO.
FT	DOMAIN	699	704	POLY-PRO.
FT	DOMAIN	814	821	POLY-PRO.
FT	DOMAIN	1340	1345	POLY-GLY.
FT	DOMAIN	1398	1403	POLY-GLY.
FT	DOMAIN	1436	1442	POLY-PRO.
FT	DOMAIN	1982	1991	POLY-PRO.
FT	DOMAIN	41	1795	4 X 57 AA TYPE A REPEATS.
FT	REPEAT	41	95	1-1.
FT	REPEAT	98	154	1-2.
FT	REPEAT	281	337	1-3.
FT	REPEAT	1740	1795	1-4.
FT	DOMAIN	337	549	2 X TYPE B REPEATS.
FT	REPEAT	337	418	2-1.
FT	REPEAT	476	549	2-2.
FT	DOMAIN	1899	2089	3 X 50 AA TYPE C REPEATS.
FT	REPEAT	1899	1948	3-1.
FT	REPEAT	1965	2014	3-2.
FT	REPEAT	2040	2089	3-3.
FT	CONFLICT	57	57	R -> A (IN REF. 2).
FT	CONFLICT	109	109	Q -> S (IN REF. 2).
FT	CONFLICT	414	414	P -> PPHRGAGNMGPP (IN REF. 2).
FT	CONFLICT	532	532	T -> K (IN REF. 2).
FT	CONFLICT	682	682	Q -> K (IN REF. 2).
FT	CONFLICT	730	730	E -> D (IN REF. 2).
FT	CONFLICT	750	750	L -> R (IN REF. 2).
FT	CONFLICT	834	834	A -> T (IN REF. 2).
FT	CONFLICT	1035	1035	G -> A (IN REF. 2).
FT	CONFLICT	1068	1068	M -> L (IN REF. 2).
FT	CONFLICT	1285	1285	P -> R (IN REF. 2).
FT	CONFLICT	1400	1400	G -> A (IN REF. 2).
FT	CONFLICT	1611	1611	T -> S (IN REF. 2).
FT	CONFLICT	1729	1729	G -> A (IN REF. 2).
SQ	SEQUENCE	2142 AA; 227840 MW; 32DDF16B9B52420A CRC64;		
Query Match 5.0%; Score 110; DB 1; Length 2142;				
Best Local Similarity 24.1%; Pred No. 14;				
Matches 55; Conservative 29; Mismatches 94; Indels 50; Gaps 11;				
QY	178	LPSPDQSPGNGATTSTRDNYHLLTEEEFGVWSQSMKWHSONKSGSVFVRGPTQEPCC--	235	
DB	516	LPAPPAPPASAPTETE-----PEEP-----AQAPPAQSTPTPGVAAAPTIV	558	
QY	236	-----SESQLKESF-VPPPTPKENKQEREDENWRLPPPPVAETVPSPSVTEIETPLQR	290	
DB	559	SGGGSTSTSSGSGFEASPVPEQLPSKEG-----PEPPEEVPPTTTPPVKVEPKGDG	610	
QY	291	IPRTATIAEPLGHCTF--TISPAFVHSLNKKRQLELLREVEWPGRGHMAATCCKLQ	348	
DB	611	IGPTROPPSQGLGYPKYQKSLPRF-----QRQQEQQLLKQOQHOHQOQOQSGSAPPTP	664	
QY	349	V-EGQDRTMSLAAPVREAPPPP-----TCASSEPSVPALP--GADPQ	388	
DB	665	VPPSPPPQVTLGAVPAPQAPPPPPKALYFALGRP--PPMPPMNFDP	710	

Search completed: December 25, 2002, 01:11:24
Job time : 32 secs